DOCUMENT RESUME

ED 392 757 SP 036 495

TITLE	Assessment and Accountability: Report from the Prichard Committee for Academic Excellence, Task Force on Improving Kentucky Schools.
INSTITUTION	Prichard Committee for Academic Excellence, Lexington, KY.
PUB DATE	Oct 95
NOTE	22p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; *Accountability; *Educational Assessment; *Educational Improvement; Educational Quality; Educational Research; Educational Testing; Elementary Secondary Education; Incentives; Public Schools; *State Standards; *Statewide Planning; Test Construction; Testing; Test Results
IDENTIFIERS	Alternative Assessment; *Kentucky; *Kentucky State Department of Education

ABSTRACT

As part of a Kentucky effort to improve public education and to increase taxpayer confidence in expanded educational funding, a citizen committee examined and made recommendations on approaches to school accountability and assessment, particularly student testing in the Kentucky Instructional Results Information System (KIRIS). The recommendations were: (1) to accelerate the creation of information that clarifies academic standards, guides the creation of questions in KIRIS, and helps teachers understand the academic content and skills that they are expected to teach; (2) to allow neither the legislature nor the Kentucky Department of Education to lower academic standards; (3) to have the Kentucky Department of Education incorporate the advances of relevant recent research into KIRIS; (4) to have the Kentucky Department of Education research and develop alternative methods and measures to supplement KIRIS with alternative recognitions of school performance as part of its regular planning process; (5) to constantly scrutinize and adjust the types and nature of school rewards and sanctions; (6) to have the Kentucky Department of Education, with public involvement, find ways to provide incentives for students and educators; (7) to find ways to make test reports useful to parents; and (8) to create an Assessment Forum to help the citizen committee review KIRIS periodically and suggest ways to continuously improve the measurement and attainment of high academic standards. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 392 757

Assessment and Accountability

Report from The Prichard Committee for Academic Excellence, Task Force on Improving Kentucky Schools

October, 1995

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R F Sipkins

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

S 1636 P95

BEST COPY AVAILABLE

Assessment and Accountability
Report from the Task Force on Improving Kentucky Schools

Holding schools accountable for student performance is a new and challenging idea. It follows that devising an effective system of school accountability and assessment is one of the most visible and controversial elements of Kentucky's new educational system. The technical and political challenges are immense. When the technical and political come together, as they do here, the challenges compound for the Prichard Committee, a group of informed citizens who are not technical experts. For perspective, we remind ourselves that Kentuckians are engaged in a serious discussion that simply did not exist before 1990 because school accountability was not even being attempted.

The stakes are high. To a large degree public acceptance of higher spending on Kentucky public schools (up 46 percent since 1989-90) was based on the promise that schools would get much better; credible testing is one of the ways to demonstrate such improvement to the taxpayers.

Kentucky is creating a new testing and accountability program to encourage students and teachers to reach higher levels of learning than ever achieved before. Accountability pushes educators and students. New tests measure school performance for accountability, and are the basis for financial rewards. Their purpose is to show the public and parents how well schools are doing at their job of educating students, and to provide appropriate consequences for schools that are effective and those that are not. The tests also are meant to drive instruction and curriculum. Because testing is driving instruction, and because it has real consequences for teachers, it is imperative that it be done extremely well and that it be

credible.

But parents expect tests, in addition to measuring school performance, to provide individual scores and national comparisons for their children. The big question is whether one test can do all of this; no one knows for certain, but most experts are doubtful. The challenge Kentucky faces is being confronted all over America, as all states attempt to create high and measurable academic performance standards. Researchers who have studied the Kentucky Instructional Results Information System (KIRIS) say that Kentucky should not go back to the standardized tests used earlier because they were damaging to good teaching. They argue instead that Kentucky must press forward in its attempt, begun with KIRIS, to develop a test that genuinely improves instruction and measures school performance. They readily acknowledge that this is hard to do, given the state of testing technology.

Taken together these conditions effecting Kentucky education—the centrality of testing to instruction and accountability, the test industry's limitations for creating tests that Kentucky needs, and the political nature of decisions about testing—provide Kentucky decision makers and Kentucky schools with a serious dilemma.

Testing to determine how much students learn was controversial even before it had consequences for teachers. Kentucky began to require a statewide test in 1979. (Some school districts used standardized tests before then.) Three different tests were adopted and abandoned between 1979 and 1990: two versions of the Comprehensive Test of Basic Skills (CTBS), and the Kentucky Essential Skills Test (KEST). Before KIRIS, critics cited numerous weaknesses in multiple choice testing such as those used in Kentucky. Dan Koretz of the Urban Institute observed that "there can be no doubt that current norm-referenced tests

overstate achievement levels in many states, often by large margins." On Kentucky's 1987-88 test for instance, students in every school district scored above the national average. Critics, such as the University of Louisville Professor George Cunningham, claimed the 1986 test was "seriously flawed." Test bias, "dumbing down," narrowing of the curriculum, score inflation, and parent confusion were common.

Clearly change was needed. New testing came to Kentucky with the passage of the Kentucky Education Reform Act. Debates over the new test have been particularly intense because, with reform, testing finally meant something—rewards and sanctions were to be assigned to schools based on student academic progress. With this decision the testing debate took on serious and new complications.

The idea behind Kentucky school reform is to set high standards and provide encouragement to teachers to reach all students—the most gifted to the least gifted—with high-level academic instruction. To do this, academic standards were to be set and measured with a new test upon which rewards (incentives) and sanctions were to be based. Forty-nine of 50 states are currently attempting to set and measure academic standards. Kentucky, however, stands apart by tying financial rewards to test scores and other measures of school effectiveness. Political pressure for quick results makes investing the time needed to develop valid and reliable tests difficult. Some researchers say that Kentucky has moved too quickly in its testing program because of that pressure, but we cannot allow that to be the basis for failure to make the best effort possible to create effective assessment.

This whole process—devising a fair and effective system, including a new test, for assessing student progress and holding schools accountable for student learning with

consequences—is immensely complex. It is the part of reform where the most divisive politics and the most thorny technical problems come together. In the end, decisions about assessment require political judgment. However, testing is also a technical challenge. And the technical expertise needed to create more authentic testing for American schools is, it appears, less potent than hoped for in the 1980s. This weakness then compounds Kentucky's challenge and makes it even more imperative that Kentucky stay the course.

Kentucky's task is difficult because it includes complex technical problems at the cutting edge of the nation's testing industry. It is also difficult because no off-the-shelf test exists, ready for classroom use. It is difficult because there is disagreement over whether financial rewards are effective incentives for encouraging employee performance, even in the business community. It is difficult because Kentucky citizens and teachers have never before tried to decide what all students should know and be able to do, a task delegated to textbook publishers, national testing corporations, and individual teachers before 1990. It is difficult because standardized tests are misunderstood and confusing to parents and the public. It is difficult because many educators object to the very concept of measuring performance and giving financial incentives on the basis of that performance. It is difficult because teachers don't all know what to do in the classroom to reach new or higher standards. And it is difficult because it has become the most politicized element of education reform; David Cohen of Michigan State University says testing is at the center of a "ferocious polemical debate."

What's To Be Done

Virtually everyone making proposals for improving American education agrees that

high academic standards are critical and that many paths are available to reach that goal. Diane Ravitch, former Assistant Secretary of Education in the Bush administration, writes:

> Is the goal—higher levels of academic achievement for all students—worth the effort? Absolutely. Although not every student will reach the highest levels of performance, all students can learn much more than they do now and improve their academic performance.
>
> Irving Louis Horowitz of Rutgers University has described the shaping of standards as 'a way of doing things by identifying or creating or constructing models of performance to which presumably rational persons can aspire.' Two points in his description bear remembering. First, identifying models of performance must be a process of continuous improvement; second, the models of performance that serve as standards must be better than common practice; they must be models to which 'presumably rational people can aspire.' 'Deep risks' must be taken in the process of setting and revising standards, Horowitz observes, but 'there are catastrophes in the failure to run such risks'."

To focus our thinking on this complex task, the Prichard Committee gathered and listened to experts in the field of testing and accountability. In June, 1995, the committee convened an assessment forum with representatives of research centers, universities, and consulting firms all recognized for their expertise. (Charlie Abelman, Harvard University Graduate School of Education; Eva Baker, Center for Research, Evaluation, Standards, and Student Testing, UCLA; Anthony Bryk, University of Chicago; Tony Cipollone, Annie E. Casey Foundation; Tom Corcoran, Consortium for Policy Research in Education; Jane David, Bay Area Research Group; Susan Fuhrman, Consortium for Policy Research in Education; Paul LeMahieu, College of Education, University of Delaware.) In addition, two serious studies of KIRIS have been released in the past six months, one by the Kentucky Institute for Education Research and the other by the Office of Education Accountability.

From the perspective of the experts assembled by the Prichard Committee:

> Kentucky's education reform (KERA) and its approach to assessment

represent the most comprehensive and boldest effort in the nation to improve public schooling. KIRIS represents a significant step in creating new, more challenging assessments that can guide schools to help all students reach world class standards. Developing such a cutting-edge assessment system is an enormous challenge and, as such, will inevitably have flaws. The issues raised by KIRIS include some of the thorniest measurement issues in American education. As Kentucky learns along the way, so does the nation.

No major assessment has ever been built so openly or with so much scrutiny. Nationally norm-referenced standardized tests were developed behind closed doors. The familiar percentiles and grade-equivalent scores, and the construction of items in these tests, receive considerable technical criticism, and had they been developed as openly as KIRIS would have raised as many questions. The Kentucky Department of Education deserves considerable credit for providing data about KIRIS, involving many people in its development, and acknowledging that there are problems. In fact, these problems were anticipated by the legislation which characterizes the assessment as under development through 1995-96. Consequently, fixing KIRIS does not represent a change in course.

Some of the problems with KIRIS are easy to fix; some are more difficult. There is no question that improving KIRIS is worth the effort and that abandoning it would severely undermine the progress of KERA. It is clearly successful in sending a powerful signal to educators, students, and parents that schools must change—that there is much more to learning than basic skills and lists of facts. The only alternative to fixing KIRIS is unacceptable—a return to traditional norm-referenced tests which have a long history of narrowing the curriculum and impeding school improvement." (Jane David to Robert Sexton, July 5, 1995, "Review Draft Summary of KIRIS Meeting.")

In this arena the challenge to citizen volunteers on the Prichard Committee, who are not technical experts, is to select and address the topics that are most important without becoming bogged down in technical or administrative detail. Our goal is to encourage vastly improved public schools for all Kentucky children. An important method to achieve that overall goal is to perfect the assessment and accountability system. Many topics are important in the assessment debate, including many raised in the reports by the Office of Education Accountability and the Kentucky Institute for Education Research, but we as

citizens need not address each of them.

Keeping our goal in mind, we make recommendations on these topics:

- Providing continuous improvement in testing and using research in an ongoing fashion.
- Making academic content more clear to teachers and parents.
- Retraining teachers to help students reach higher standards.
- Varying and expanding the measures that are used to evaluate schools.
- Finding effective combinations of rewards and sanctions.
- Devising effective interventions in schools so they will improve.
- Providing incentives that encourage students to perform well.
- Providing individual student scores and comparisons between Kentucky students and students in other states.
- Helping parents understand the limitations of testing and the nature of absolute academic standards.
- Confronting the political challenges of accountability.

Recommendations

Underlying all these recommendations is our commitment to the belief that public schools should demonstrate, in clear and visible ways to parents and taxpayers, what they are contributing to children's learning. Difficulties and the time required notwithstanding, we believe the pursuit of and commitment to school accountability is imperative and possible for improving the quality of education for all Kentucky children. We believe that the citizens of Kentucky, by their support of increased spending for public education in the 1990 legislative session, expected improvement in education quality. Determining fair and understandable ways to show that improvement to the public is difficult and will take time, as seen in the

events surrounding standardized testing over the past 15 years.

I. We recommend *that the Kentucky Department of Education accelerate the creation of information that clarifies academic standards, guides the creation of questions in KIRIS, and helps teachers understand the academic content and skills that they are expected to teach. This information may include curriculum frameworks, content standards, and core concepts that are basic for all students. This process should be open and public, engaging all who express interest, but should be primarily the responsibility of teachers. This should be done recognizing the need for balance between local and state authority. We recognize the difficulty as well as the need for this balance. There is a tendency for state policies and guidelines to be resented and criticized as mandates and for local authorities to be considered autonomous and absolute. Neither tendency is acceptable. Coordinated and reasonable consistency of school curriculum requires statewide policies and guidelines be developed with care and with latitude for local differences and initiative.*

We also recommend that the academic expectations created by teachers, principals, parents, and university professors in 1991 and then revised in 1993, be re-examined on a regular cycle, every four years.

Rationale

Testing is influencing what happens in classrooms across Kentucky in positive ways. Students are doing more writing, more explanation, and more hands-on activities. However, there is concern that KIRIS, combined with directives from the Kentucky Department of Education, may have swung teaching too far away from basic skills and content knowledge

toward an emphasis on problem solving and application of skills. Ongoing research to verify the influence of KIRIS on classroom practices will be important.

The challenge for the state is deciding how to inform teachers about what they are expected to teach in ways that ensure they are adequately preparing their students for KIRIS without overly constraining school choices about curriculum. What is the right balance of basic skills, content knowledge, application of skills, and problem solving, for example? What is the best form for communicating curricular guidance that is neither too general to be useful (short lists of big ideas) or too long to be usable (detailed lists of everything)?

There is no easy solution to this problem. Every state faces this issue of what appropriate curriculum content is and who determines it. Until now these decisions have mostly been left up to the textbook companies and publishers of standardized tests. Teachers need guidance that falls somewhere between that contained in curriculum frameworks and content guides. The guidance needs to be supported by professional development. As more KIRIS items with examples of student work are released each year, teachers will better understand what they are expected to do.

Two special challenges have become clear since 1990. Citizens and parents have never before been engaged in the process of setting curricular standards. They left those decisions to textbook publishers and standardized test makers, and those decisions reflected national needs and markets, not local needs. Because setting curriculum is terribly important and reflects basic community values, it can be divisive and difficult. (This is probably one reason that parents and citizens were not engaged in the process before.) This divisiveness can, if permitted, disrupt all attempts to improve school quality and this must be avoided.

Improving the quality of education for all students is a more important goal than adult debates over ideological or political issues. These debates should not disrupt the education of children.

A second complicating factor is that the academic standards that are tested serve as important guides to teachers. Teachers, attempting to carry out the mandates given to them, want and need clear direction; in the worst case they ask to be told, in detail, what to teach. When academic expectations are changed too often, it disrupts teachers' abilities to meet expectations and, in the extreme, provides an excuse to do nothing.

To see that students are educated well by teachers and that instruction not be disrupted, we recommend that Kentucky's Academic Expectations be re-examined on a regular schedule, not haphazardly or for political whim, with teachers themselves most heavily involved.

II. We recommend *that neither the legislature nor the Kentucky Department of Education lower academic standards. Instead, all energies should be concentrated on helping teachers and students achieve the high standards that have been established. To improve public and parental understanding of these standards, the Kentucky Department of Education should find additional ways to report student achievement to parents and, if technically feasible, provide more levels of achievement in reports to parents and the public. This reporting should identify achievement in specific basic skills in addition to those skills and knowledge that go well beyond the basics.*

Rationale

The political impulse is to lower standards if high standards are difficult to achieve. In Kentucky there are signs of this impulse to lower standards in suggestions about replacing academic standards with required curriculum and adopting a machine scorable multiple choice test to replace performance testing.

The Prichard Committee understood the power of required courses when it recommended in 1981 that a precollege curriculum and admission standards for public universities be established.

Requiring courses, however beneficial, is not the same as requiring standards of achievement and is de facto lowering of academic standards. Academic standards and required courses are not the same. Academic courses (or content requirements) are meant to ensure that a student is exposed to certain material, knowledge, or skills. Academic standards, on the other hand, are designed to ensure that the student learns, masters, or demonstrates competency. A metaphor occurs in athletics: in teaching a child to swim the instructor both shows the student what swimming looks like (content) and requires that the student stay above water (performance standard). In the past, schools measured only content.

In a state like Kentucky, with historic educational deficiencies, a return to academic mediocrity or worse is totally unacceptable.

There is, however, much confusion about the standards that have been set. Standards are not seen as valid when they contradict what people know to be true. For example, the standards do not make sense to people if students who graduate and receive college scholarships score "novice" (assuming they have put out effort on the assessment). On the other hand, if the standards are judged to be world class, they may not be too high

regardless of public perception.

The reasonable appearance of standards is also a function of the confidence in the process that produced them, and of how they are measured, not simply what they say. Because the standards are high and their achievement is expected to take 20 years (and assumed to move in equal steps over that time), there should be consideration of more frequent and more attainable standards and perhaps more reporting levels.

III. We recommend *that the Kentucky Department of Education proceed by February, 1996, to make adjustments and demonstrate to the General Assembly and the public that the concerns raised in recent research have been successfully addressed. It is imperative that the basis for rewards and sanctions be reliable and valid when held up to the scrutiny of researchers and the public. Significant changes in the KIRIS assessment will be required, as identified in research. If the above is accomplished, there should be no need to delay or alter the schedule of rewards and sanctions. This recommendation is also based on the continued and on-going responsibility of the Department to employ the best technology available as research advances the capability for testing. We believe delay awaiting technical advances would seriously impair the opportunity to fully evaluate the advantages of improving educational quality based upon an incentive program.*

Rationale

The dilemma in the decision about whether to delay rewards and sanctions is that the Commonwealth of Kentucky has made a commitment to educators to reward good performance. On the other hand, the educational community, including state officials, has

made a commitment to the public that improved performance will be demonstrated in an understandable way.

The two reports recently published (by the Kentucky Institute for Education Research and the Office of Education Accountability) argue that KIRIS is currently not reliable enough to serve as a basis for rewards and sanctions but that it can be made so by expanding the types of test items and by coupling KIRIS with additional information about schools for purposes of determining rewards and sanctions. Adding multiple choice questions and other types of machine scorable items and reducing the weights assigned to the least reliable elements will also increase reliability, these researchers argue.

However, it is also clear that the research community is deeply divided on these matters and citizens should be wary. Despite criticisms, the researchers cited above also say that Kentucky's test is the best effort yet devised to measure student performance and that no other "off-the-shelf" test exists for purchase that is adequate to meet Kentucky's needs.

The decision to make the recommendation above was reached before the Response to the OEA Panel Report (Edward H. Haertel and David E. Wiley, September 19, 1995) was available, but that response is worthy of note because it questions some of the emphasis and recommendations in previous research by the Office of Education Accountability. In particular, it states:

> We consider both inappropriate and irresponsible the report's allegation that KIRIS is seriously flawed and needs to be substantially revised and that the public is being misinformed about the extent to which student achievement has improved statewide . . . . (p.1).
>
> In Chapter 8, the panel offers a thoughtful and, on the whole, cautious review of the evidence available from other sources concerning changes in student achievement in Kentucky. They point out, and we concur, that changes

> in motivations, teaching to the test, increased familiarity of students and teachers with novel assessment formats, and outright cheating may all have contributed to measured improvement in KIRIS scores, but the relative magnitude of their contributions versus real changes in student proficiency are unknown . . . . (p. 7).

These and other comments critical of research on KIRIS and their recommendations, underscore our thinking that continued improvement in the test technology is essential but also that the benefits of motivation, which appear to be generated by the accountability in education reform, should not be jeopardized by delay.

Kentucky's testing system must be made as valid and reliable as possible because it is essential to widespread improvement in student learning as evidence for rewards and sanctions and as the linchpin for the 1990 political agreement to increase school funding through increased taxes. Researchers have suggested numerous ways that the test can be improved. The Kentucky Department of Education should proceed to make these adjustments and present them to the 1996 General Assembly.

IV. We recommend *that the Kentucky Department of Education proceed as soon as possible with research and development of alternative methods and measures to supplement KIRIS and with alternative recognitions of school performance as part of its regular planning process.*

Rationale

It is suggested in the research on KIRIS that performance testing alone will not adequately measure school performance. States around the nation are examining a variety of other measures. Kentucky should join in that process.

Decisions about rewards and sanctions might be based on additional and more in-depth information about schools that are directly related to student performance, such as features of the school's curriculum, the materials used, and the kind of work assigned to students. These might increase the emphasis on results such as higher learning or work, attendance, retention, and transition to postsecondary education or a job or add new non-cognitive measures. It might also include measuring the implementation of other components of good instructional practice that influence student performance such as the primary program, extended school services, and family resource centers. Expanding the information base for accountability lessens the reliance on KIRIS as the primary determinant. This, in turn, lessens the pressure for unassailable accuracy. By increasing the information base and the methods by which school progress is determined, the motivation and opportunity to cheat on the tests is reduced.

These types of information could be obtained by school quality review or inspectorate teams, similar to those in some other states and countries. These will provide a fuller picture of what schools are doing, both for purposes of accountability and for providing information to schools on areas needing improvement. Some could also be gathered through surveys to researchers and to students. Different ways of defining and collecting this information should be tested on a pilot basis.

However, under no circumstances should the provision of assistance be delayed. There are only advantages to providing assistance to poorly performing schools, even if their performance has been inaccurately measured. All schools, but especially those with the poorest performance, can benefit from assistance and professional development opportunities. These should continue under any scenario.

V. We recommend *that the types and nature of school rewards and sanctions be constantly scrutinized and that adjustment be made as needed.*

Rationale

There are several questions about the current incentive structure that is part of the accountability system. One is whether the rewards and sanctions as currently defined operate as effective incentives for teachers to improve their teaching. Other questions arise about practices such as the way rewards are distributed within schools. For instance, should rewards go to schools or teachers, and what are the consequences of these decisions? Is it possible, for instance, that giving financial rewards to schools and not to teachers would be more popular with teachers and the public? Would changing this arrangement harm instruction? There are no clear answers to these important questions.

It is also possible that the formula for granting rewards may sometimes have negative consequences. Research on high school restructuring (Fischetti, et. al., 1995) suggests, for example, that the heavy weight given to KIRIS as opposed to student retention encourages high schools to push students out of school. These consequences should be constantly monitored.

VI. We recommend *that the Kentucky Department of Education, with widespread public involvement, devise methods for providing incentives for students as well as educators.*

Rationale

There is concern, particularly among teachers, that there should be incentives for students (and/or their parents) as well as for teachers. In the national discussion of standards,

incentives for students are being aggressively promoted by the American Federation of Teachers. The danger in our view, is that providing meaningful and fair incentives is easier to talk about than to do. For instance, incentives are totally different for students in early grades than for those ready to graduate from high school. Many advocates of student incentives gloss over these differences.

However, consequences for students make sense when there is input from parents and when attention is paid to whether students have had the opportunity to learn what they are expected to know. Several states have adopted examinations required for high school graduation and these, although not without drawbacks, should be considered. Ultimately, it is employers, institutions of higher education, and parents who control real consequences for students. Communities and families must provide the most meaningful incentives for students; if the community and employers don't value learning, why should students? Consideration might be given to partnerships with parents to criteria for graduation tied to KIRIS and perhaps to criteria for earlier transitions, such as primary to intermediate and intermediate to middle school.

VII. We recommend *that the Department of Education find ways to make test reports more useful to parents while being straightforward about what the KIRIS test, or any test, can and cannot do. This recommendation suggests that some improved multiple choice questions be combined with or added to KIRIS so that individual scores and some measure of national comparisons can be provided to parents, and that the General Assembly provide for the increased costs of such testing in the education budget.*

We also recommend that schools create their own ways to report regularly, clearly, and openly on student learning to parents and the public. This reporting should emphasize student work, not test scores.

Under no circumstances, however, should Kentucky return to an examination that is totally machine scorable multiple choice and not based upon student demonstrations of high quality academic work.

Rationale

We make this recommendation because many parents express the desire for measures that compare their children with other children. Until a performance-based testing system in which parents have confidence and which measures achievement against an absolute standard is created, multiple choice questions will be useful. (About 70 of Kentucky's school districts use a multiple choice, machine scorable assessment, in addition to KIRIS, at this time.)

We make this recommendation mindful of widespread agreement in education research that traditional standardized testing has contributed substantially to the current problems in American education. Since such testing is damaging to student learning it should be used sparingly. Indeed, at least two researchers cited recently as critical of KIRIS were also critical of standardized testing in the 1980s, using similar language to criticize both Kentucky's old and new tests. Weaknesses in the fields of testing and psychometrics make reform efforts in states like Kentucky particularly difficult.

KIRIS results can be made more useful, but it is important to be prudent and conservative about what KIRIS can and cannot do. It cannot be all things to all people. No single test can serve all purposes including school accountability, guidance to teachers,

diagnostic information on individual students, and results on progress for parents. It is important to clearly communicate which purposes KIRIS is designed to serve. Otherwise, unmet expectations will undermine its credibility.

Usefulness of reports is connected to what is included on the assessment. For example, items that assess basic skills and that can be reported separately would be viewed as useful by many. Similarly, reports will be perceived as more useful if they incorporate national norms, such as percentiles, for example, by including test items that have been norm referenced.

Usefulness of results also depends on when the testing occurs. The choice of grades 4, 8, and 12 for the first developmental phase of KIRIS, chosen in part to mesh with the National Assessment of Educational Progress (NAEP), may make less sense than testing that matches the organization of schools. For example, testing at the end of the primary grades (3), intermediate grades (5), middle school (8), and grade 11 may be more useful.

Results also need to be reported in ways that communicate what is expected. Results must tell teachers whether their curriculum and instruction is on track and, if not, what they need to do differently. Educators must be able to evaluate the information they get back from the testing if they are to improve their practice. Teacher training and scoring are important parts of understanding what it takes to produce high quality student work.

VIII. We recommend *that the Prichard Committee create an Assessment Forum to help the Committee's citizen volunteers review KIRIS periodically and to suggest ways to continuously improve the measurement and attainment of high academic standards for all Kentucky*

children. The Assessment Forum will be composed of national experts (such as those already assembled by the committee) as well as the public and educators. The Forum should assist the Committee with its review of rewards and sanctions (Recommendation I) alternative measurement (Recommendation IV), improvements to KIRIS (Recommendation III) and student incentives (Recommendation VI).

Rationale

This forum will make available to the committee and to other Kentucky citizens the expertise needed to solve the difficult issues we have identified. It will give the volunteer members of the Committee, who are not technical experts on testing, the capacity to examine new technical issues as they arise. It also will serve as a safe space for the public and parents to express their concerns and propose alternatives.